

Web Security using Transparent Image Captcha (TIC)

Ved Prakash Singh, Preet Pal, Pushpendra Kumar Pateriya

*School of Computer Science, Lovely Professional University
Phagwara, Punjab*

Abstract— since the evolution of Internet, web security has become a primary concern and is also very challenging with its day by day growth. Our aim is to utilize CAPTCHAs as a system to improve the security for Internet based applications. There are many threats circulating on the Internet and one such threat is BOT, a malicious program, which runs automated tasks over the Internet and thus creates problem. As a shield to this problem we make use of CAPTCHA. The main idea behind this research is to provide a better methodology implementing the web security. We will try to develop the technique which will use the transparent image based CAPTCHA. Along with this we will also use the filters to increase the complexity in identifying the images. This will help in preventing the attacks from the Internet like BOT etc. The transparent images based CAPTCHA system makes exploitation a very difficult thing for BOT, but it is easier for a human being to interpret it.

Index Terms—CAPTCHA, BOT, Turing Test, Security, HIP,OCR.

I. INTRODUCTION

First time CAPTCHA was invented in 2000 at Carnegie Mellon University by John Langford, Nicholas J. Hooper and Luis Von Ahn [1], [7]. CAPTCHA is an acronym for “Completely Automated Public Turning Test to tell Computers and Humans Apart” [2]. The progress of Internet, Web security has become an important issue. There are too many malicious threats across the Internet which may compromise your system in the absence of any secure application which provides protection against such threats. One such threat is the Bot. A Bot is a malicious program which has the capability to run automated tasks over the network and thus creating problem in the network [3]. CAPTCHA is one such shield which can be used as a protection from these malicious programs like Bot.

A CAPTCHA may come in various forms like text based or image based CAPTCHA. The Bot operation is similar to reverse “TURING TEST” (given by Alan Turing) [2] where the program acts like judge and the other person acts like user. If the user fails this test then he/she is considered to be a machine otherwise the user is considered to be an authentic user or a human being. CAPTCHA is a defensive system that acts as a tool to check web Bots from exploiting online services on the internet including free email providers, wikis, blogs etc. It is a HIP system that is widely used to secure the internet based applications. It is also called as a challenge response test which gives a challenge to the users, when the user gives correct answer he is considered as human otherwise a web bot. CAPTCHA

is an authentication process based on challenge-response authentication. CAPTCHA provides a mechanism with the help of which a user’s can protect themselves for spam and password decryption by taking a simple test. In this test a user will see either an image or a text which are normally distorted. The user is supposed to enter the pattern exactly as shown to him if the CAPTCH is based on text. If the CAPTCHA is based on image the user is supposed to enter the correct name of the image which correctly symbolizes the image. CAPTCHA is used where authenticated access is the primary concern. Various web services like Yahoo, Google, and Bing etc. use CAPTCHA to differentiate between an authenticated user and a malicious program. CAPTCHAs are also used in the sites which provide access to sensitive data, such credit card accounts and banks.

II. CATEGORIES OF CAPTCHA

CAPTHCAs means presenting a challenge response test to the users or humans. They are classified based on what is distorted that is whether characters, digits, or images. Some types of CAPTCHAs are given below [7] :

1. CAPTCHAs based on text.
2. CAPTCHAs based on image.
3. CAPTCHAs based on audio.
4. CAPTCHAs based on video.
5. CAPTCHAs based on puzzle.

1) CAPTCHAs based on text: Text based CAPTCHAs is a very simple to implement. In text based CAPTCHA asks to end users to recognize the word that has been represented in a distortion way. And this type of CAPTCHA is natural to end users. We can also say that text based is a very easy to use without any learning type institution like as learning and training.it most widely used in many famous websites like as Gmail, You Tube, PayPal, Hotmail, Facebook, Yahoo and etc. It is secure to defined automated program if they are properly designed like as distorted form of a word that is not recognized by robots easily. Yet if the word based CAPTCHA is misrepresentative then it hard to recognize to end users. It is very effective and requires a large question bank. In Text based captcha the Number of classes of characters and digits are very small so the problem occur for user to identify the correct characters and digits. The text based captcha is possible to identify the character and digit through Optical character recognition (OCR) technique [2], [4]. In Text based CAPTCHAs simple asked questions like as based on arithmetic equation some example are given below:

- A) What is three plus two (3+2=?).
 - B) What is six minus one (6-1=?).
 - C) Which of cabbage, apple and table is vegetable?
- Example of text based CAPTCHA.



Fig 1: Text Based Captcha

2) CAPTCHAs based on image: In Images based CAPTCHAs, end users have to identify the focus of an image .And images based CAPTCHA usually interacts with users by using a mouse .Graphics-based CAPTCHAs are challenge-tests in which the users have to guess those images that have some similarity. For example: visual puzzles. In image based CAPTCHAs user is required to identity image. The advantage of image based CAPTCHA is that pattern recognition is hard AI problem and therefore it is difficult to break this test using pattern recognition technique.



Fig 2 Images Based Captcha

3) CAPTCHA based on audio: Audio-based CAPTCHAs are based on the sound-based systems. These CAPTCHAs are developed for visually disabled users. It contains downloadable audio-clips. In this type of CAPTCHA, first the user listens and after that submits the spoken word [2]. The first sound-based system name ECO was implemented by the Nancy Chan a student from the City University in Hong Kong. The audio-based system is based on the difference in the ability between computer machines and humans in recognizing spoken language. The program chooses a sequence of digits and words randomly and renders the words and number digits into sound clips and distorts it. The distorted sound clip is then presented to the user to enter the right word or number. The user is asked to enter exactly the same words as spoken the audio clip [6].

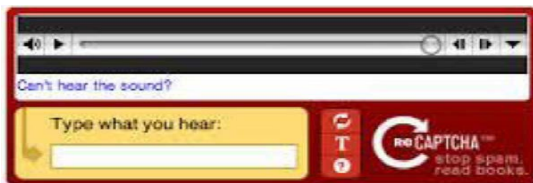


Fig 3 Audio Based Captcha

4) CAPTCHA based on video: Video CAPTCHA is a newer and less commonly seen CAPTCHA system. In video-based CAPTCHAs, three words (tags) are provided to the user which describes a video. The user's tag must

match to a set of automatically generated ground truth tags then only the test is said to be passed. The term video CAPTCHA is used to any CAPTCHA that uses a video as its means to present information to a user [2,6]. Although video CAPTCHA is limited, both commercial and academic application do exist.



Fig 4 Video Based Captcha

5) CAPTCHA based on puzzle: Usually in puzzle based CAPTCHA a given picture is divided to chunks [2], 5]. A user is supposed to combine these chunks so as to form the complete picture same as the original one.



Fig 5 Puzzle Based Captcha

There are number of applications of CAPTCHA on the web which are defined as follows:

1. Registering the web forms: There are many sites on the Internet which provide free registration to avail their services. But they are susceptible to web bots. It may come into the form of scripts which can register thousands of email accounts on the internet, thus wasting the precious space of web.
2. Online polling sites: These sites takes user's response or feedback in the form of questionnaires. To ensure that only human makes the response they make use of CAPTCHA.
3. Free Email Services. Numerous companies like as Yahoo!, Microsoft, etc. they offer free email services ,most of which they suffer from a specific type of attack like as bots that can be sign up for and thousands of email accounts a few minute or seconds. For this purpose and handle this situation that it can be improve by requiring users to prove they are human before they get free create a mail account. For instance Yahoo! Uses a CAPTCHA of to prevent bots from registering for accounts. And this CAPTCHA asks to users to read a distorted word such as the one shown below in which computer program are not as good as human to reading this types of text.



Fig.1.6 The Yahoo! CAPTCHA

4. Search Engine Bots: In search engine bots some web sites do not want to be indexing by search engine due to an html tag to prevent search engine bots from reading web pages and the tag does not guarantee that bot would not read the web page resources. The large companies used search engine bots to prevent no such kind of unwanted bots, thread, worm, spam, access the data and resource.
5. Worms and Spam: some CAPTCHA also used to prevents email spam and worms.in such kind of CAPTCHA they facility that only accept email if we know there is a human behind to the computer.
6. Preventing Dictionary Attacks: In preventing Dictionary attack CAPTCHA is used to prevent dictionary attack in password system. To prevent a computer from being too able to iterate through the entire space of password by requiring a human to type password.

III. PROPOSED ALGORITHM

Internet is very large network which comprises of many people over the network who just for their fun use it in an inappropriate way by using malicious software. In most of the cases these malicious attempts destroy vital resources in the network. To prevent the network from these threats various types of security mechanisms are used. CAPTCHA is one such security mechanism which gives user the privilege to protect their vital resources when exposed in the network. After text based CAPTCHA, Image based CAPTCHA was introduced, which require each users to identify different type rotate images or labeled type the image. That is evinces a larger gap between Bots compared to the text-based ones and human users, because of the bad ability of bots in obtaining characteristics of images. We know that a lot of web services including GOOGLE use CAPTCHA to help remove unauthorized account entry. In this image based captcha, I am going to make it as multi filtrations of the images to remove the system hacking from the BOT. After filtering the given images, will be changed. Only the human beings can able to identify that which one is the given image. And selecting the right image, the machine will match whatever that has been picked up by the user is correct or not.

Transparent image based CAPTCHA system is idea to keeping in view that Turning test is perceivable by human but is very difficult to crack by BOT. It is based on the following steps:

- 1) First of all, we have to collect the images from various sources such as internet or any other sources and store them in the database.
- 2) Now we display the images which we have collected in the grid which is 3*2 size from the database.
- 3) Then we have to randomly generate an image number between one and six and based on this random generated value the system placed an image to the grid.
- 4) Next, the user is required to choose the correct both image in sequences order only from the grid.. If the image is found to be correct, only then the system will proceed.

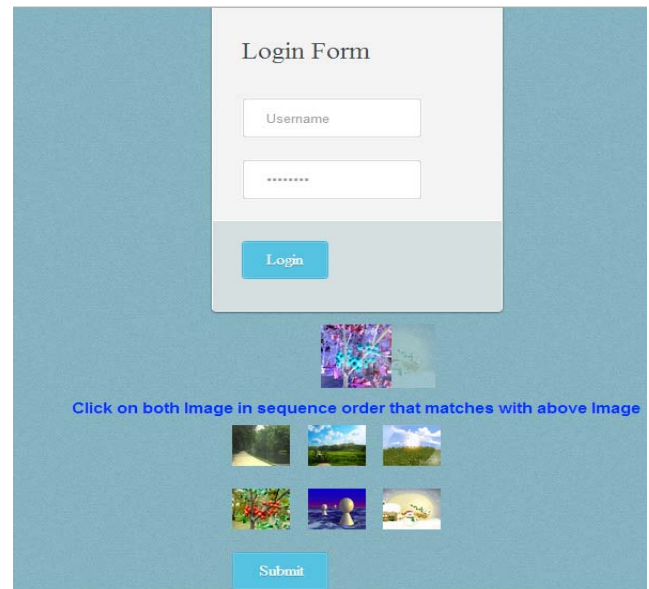


Fig 6:User interface

Flow chart :

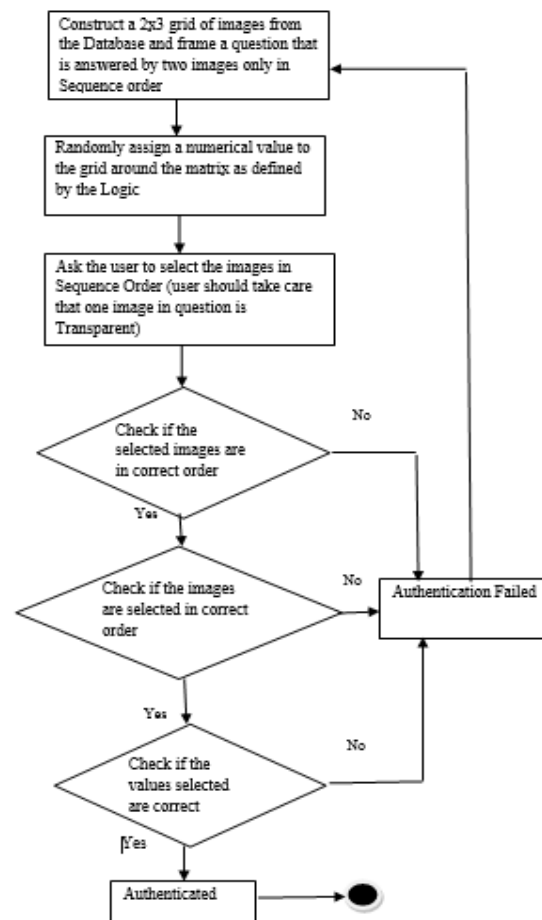


Fig 7: Flow chart of proposed algorithm

CONCLUSION

The above stated algorithm is one of the best of its kind as it provides the ease of access to the user and the security level is high. We will now start to implement this algorithm and after the completion of the task, we will get a system which is capable of preventing the BOT attacks by providing the extended security feature through this algorithm. We will also focus on the execution time of this algorithm. Our prime motive will be to reduce the execution time and provide better security.

REFERENCES

- [1] Wei-Bin Lee, Che-Wei Fan ,Kevin Ho, Chyi-Ren Dow , and "A CAPTCHA with Tips Related to Alphabets Upper or Lower Case," in Seventh International Conference on Broadband, Communication, Wireless Computing and Applications, 2012.
 - [2] Baljit Singh Saini and Anju Bala "A Review of Bot Protection using CAPTCHA for Web Security," IOSR Journal of Computer Engineering, 2013, pp. 36-42, 2013.
 - [3] Xiao Ling-Zi and ZHANG Yi-Chun "A Case Study of Text-Based CAPTCHA Attacks," in International Conference on Cyber-Enabled Distributed Computing and Knowledge Discover, 2012.
 - [4] Chen-Chiung Hsieh and Zong-Yu Wu "Anti-SIFT Images Based CAPTCHA Using Versatile," IEEE, 2013.
 - [5] Rich Gossweiler, Maryam Kamvar and Shumeet Baluja "What's Up CAPTCHA?A CAPTCHA Based on Image Orientation" WWW 2009 MADRID!, pp. 841-850, 2009..
 - [6] Wenjun Zhang, "Zhang's CAPTCHA Architecture Based on Intelligent Interaction via RIA," Research Institute of Applied Computer Technology, IEEE, 2010.
- Ved Prakash Singh and Preet Pal "Survey of different types of CAPTCHA",International Journal of Computer Science and Information technologies, 2014, pp.2242-2245.